JOURNAL OF BACTERIOLOGY, May 1990, p. 2259–2266
0021-9193/90/052259-08$02.00/0

Vol. 172, No. 5

Regulation of the *Salmonella typhimurium aroF* gene in *Escherichia coli*†

GLORIA KRESSIN MUDAY AND KLAUS M. HERRMANN*
Department of Biochemistry, Purdue University, West Lafayette, Indiana 47907–6799

Received 23 October 1989/Accepted 20 January 1990

The *Salmonella typhimurium aroF* gene, encoding the tyrosine-sensitive 3-deoxy-D-*arabino*-heptulosonate 7-phosphate (DAHP) synthase, was localized to a chromosomal *PstI* fragment by Southern blotting with an *Escherichia coli aroF* probe. This fragment was cloned by screening a plasmid library for complementation of an *E. coli aroF* mutant. The nucleotide sequence of *S. typhimurium aroF* was determined and compared with its *E. coli* homolog. The nucleotide sequences are 85.1% identical, and the corresponding amino acid sequences are 96.1% identical. The *E. coli* genes encoding the three DAHP synthase isoenzymes are evolutionarily more distant from one another than are the homologous *aroF* genes of *E. coli* and *S. typhimurium*. The *S. typhimurium aroF* regulatory region contains three imperfect palindromes, two upstream of the promoter and one overlapping the promoter, that are nearly identical to operators *aroFo1*, *aroFo2*, and TyrR box 1 of *E. coli*. The *aroFo1* and *aroFo2* sequences of the two organisms are each separated by three turns of the DNA helix with no sequence similarity. The 5′ ends of the *aroF* transcripts for both organisms contain untranslated regions with potential stem-loop structures. Translational fusions of the *aroF* regulatory regions to *lacZ* were constructed and then introduced in single copy into the *E. coli* chromosome. β-Galactosidase assays for *tyrR*-mediated regulation of *aroF-lacZ* expression revealed that the *E. coli* TyrR repressor apparently recognizes the operators of both organisms with about equal efficiency.

In bacteria and plants, aromatic amino acid biosynthesis proceeds by the common aromatic or shikimate pathway, which delivers chorismate to the terminal pathways to generate phenylalanine, tyrosine, and tryptophan (18, 32). In *Escherichia coli*, carbon flow through the shikimate pathway is controlled by modulation of the first enzyme, the 3-deoxy-D-*arabino*-heptulosonate 7-phosphate (DAHP) synthase (EC 4.1.2.15) (31). This enzyme catalyzes the condensation of phosphoenolpyruvate and erythrose 4-phosphate to DAHP (41).

In *E. coli* and *Salmonella typhimurium* there are three DAHP synthase isoenzymes that can be distinguished by their regulatory properties. The tyrosine-, phenylalanine-, and tryptophan-sensitive isoenzymes are encoded by the unlinked genes *aroF*, *aroG*, and *aroH*, respectively (11, 32). The nucleotide sequences of the *E. coli aroF*, *aroG*, and *aroH* genes have been determined (9, 33, 38). Expression of *aroF* and *aroG* is repressed by the *tyrR* gene product, the Tyr repressor, complexed to tyrosine or phenylalanine, respectively. Regulatory mutants with lesions linked to *aroF* were isolated (7, 12). Nucleotide sequence analysis localized the lesions to three operator boxes, designated *aroFo1* and *aroFo2* (12) and TyrR box 1 (7). These boxes are 20-base-pair (bp) imperfect palindromes; *aroFo1* and *aroFo2* are of similar sequences, located upstream of the *aroF* promoter, and are separated by three turns of the DNA helix (12). The TyrR box 1 overlaps the promoter and shows some sequence similarity with *aroFo1* and *aroFo2* (10). The regulatory region of *aroG* features only one operator box, which is similar to *aroFo1* and *aroFo2* (12).

In wild-type *E. coli* grown in minimal media, the phenylalanine-sensitive DAHP synthase contributes 80% and the tyrosine-sensitive DAHP synthase contributes 20% of the total enzyme activity (42). However, the *aroF* promoter scores much higher than the *aroG* promoter on a scale of relative promoter strength (12, 15), a fact that may only be reflected in *tyrR* strains. Thus, repressor binding to the *aroG* operator may be less efficient than to the *aroF* operator. In wild-type *S. typhimurium* grown in minimal media, the phenylalanine-sensitive DAHP synthase may also be the major isoenzyme (20), although the tyrosine-sensitive isoenzyme has been reported to predominate under these conditions (22). A comparison of the nucleotide sequences for the *aroF* genes of *E. coli* and *S. typhimurium* was expected to offer some insight into an explanation for the potentially different expression levels. This report describes the cloning and characterization of *aroF* from *S. typhimurium*, the construction of *E. coli* and *S. typhimurium aroF-lacZ* translational fusions, and studies on the regulation of *S. typhimurium aroF* in *E. coli*.

MATERIALS AND METHODS

Bacteria, plasmids, and bacteriophages. Bacteria, plasmids, and bacteriophages used in this study are listed in Table 1. *E. coli* GKM41 (*aroF hsdR4*) was derived from JM105 (47) in two steps by using phage P1 grown on NK5161 (*tyrA*::Tn*10*) (23) and YS482 (*aroF*), respectively. Strain YS482, isolated by R. L. Somerville, is a one-step UV-induced mutant of wild-type strain W1485. Strain YS482 lacks tyrosine-sensitive DAHP synthase. Strain GKM41 grows on minimal salts medium, but not on such medium supplemented with phenylalanine and tryptophan. Strain SP564, a CSH63 derivative obtained by transduction with phage P1 grown on PLK 1336, has a Tn*10* insertion near *tyrR* but is Tet^r and TyrR$^+$. Strain SP564-1, an SP564 derivative from which the Tn*10* has been excised by the method of Bochner et al. (4), is Tet^s and TyrR$^-$. Bacteria were grown in either Luria broth, 2× YT (29), or minimal salts medium supplemented with vitamin B_1 and biotin (45). Cells were plated on nutrient agar or M9 agar supplemented with

* Corresponding author.
† Paper no. 12209 of the Purdue University Agricultural Experiment Station.

TABLE 1. Bacterial strains, plasmids, and phages used in this study

Designation	Relevant genotype and genes	Source or reference
S. typhimurium		
LT2	Wild type	R. L. Somerville
E. coli		
YS482	*aroF*	R. L. Somerville
JM101		28
JM105	*hsdR4*	47
D1210	*lacI*^{q}	24
NK5161	*tyrA*::Tn*10*	N. Kleckner via R. L. Somerville
GKM41	*aroF hsdR4*	This study
MC4100	Δ(*argF-lac*)*U169*	40
CSH63	Δ(*lac-pro*)	29
PLK1336	*zci-223*::Tn*10* near min 28	P. L. Kuempel via R. L. Somerville
SP564	Δ(*lac-pro*) *zci-233*::Tn*10*	R. L. Somerville
SP564-1	Δ(*lac-pro*) *tyrR*	R. L. Somerville
Plasmids		
pKGW	*E. coli* endonuclease	24
pMLB1034	Promoterless *lacZ*	40
pCG201	*E. coli aroF*	12
pGM59	*S. typhimurium aroF*	This study
Bacteriophages		
M13mp18, M13mp19	*lacα*	28
M13GME	*E. coli aroF* regulatory region	This study
M13GMS	*S. typhimurium aroF* regulatory region	This study
λRZ5	Promoterless *'lacZ,' bla*	R. Zagursky via H. Zalkin
λGME4	*E. coli aroF-lacZ* fusion	This study
λGMS4	*S. typhimurium aroF-lacZ* fusion	This study

vitamin B_1 and biotin (29) and, where indicated, with 20 μg of amino acids per ml, 50 μg of ampicillin per ml, 20 μg of kanamycin per ml, and 15 μg of tetracycline per ml. M13 phages were plated in Luria broth top agar containing 40 μg of 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) per ml and 28 μg of isopropyl-β-D-thiogalactopyranoside (IPTG) per ml. λ phages were plated in tryptone top agar on minimal plates supplemented with 40 μg of X-Gal per ml.

Chemicals. Phosphoenolpyruvate (5) and erythrose 4-phosphate (39) were synthesized and assayed as described previously. Kanamycin, chloramphenicol, tetracycline, tyrosine, phenylalanine, and tryptophan were from Sigma Chemical Co.; nutrient agar was from Difco Laboratories; agarose and IPTG were from Bethesda Research Laboratories, Inc.; X-Gal, dideoxynucleotides, and 7-deaza-dGTP were from Boehringer Mannheim Biochemicals; [α-^{32}P] dCTP was from ICN Radiochemicals or Amersham Corp.; low-melting-point agarose was from FMC Corp., Marine Colloids Div.; redistilled phenol and 0.2-mm wedge-shaped sequencing gel spacers were from International Biochemical, Inc.; deoxynucleotides were from Pharmacia, Inc.; acrylamide, hydroxylapatite-HTP, and protein molecular weight markers were from Bio-Rad Laboratories; and formamide and propanediol were from Eastman Kodak Co. Oligonucleotides were synthesized by the Purdue Laboratory for Macromolecular Structure.

Enzymes. Restriction enzymes and T4 DNA ligase were from Bethesda Research Laboratories or New England BioLabs, Inc.; restriction enzyme and ligase buffers and exonuclease III were from Bethesda Research Laboratories; Klenow fragment of DNA polymerase I was from Boehringer Mannheim Biochemicals or New England BioLabs; Klenow sequencing kit, T4 polynucleotide kinase, and mung bean nuclease were from New England BioLabs; Sequenase (T7 DNA polymerase) and Sequenase sequencing kit were from U.S. Biochemicals Corp. DAHP synthase was assayed as described previously (37).

DNA manipulations. Chromosomal DNA was isolated by the method of Saito and Miura (34); agarose gel electrophoresis, ligations, plasmid and phage miniscreens, electroelutions, and transformations were as described previously (1).

DNA probes were labeled by nick translation with [α-^{32}P] dCTP. Genome blotting to dried agarose gels (44) was performed under optimal hybridization conditions (2, 27). *S. typhimurium* DNA, digested with restriction endonuclease *PstI* and size fractionated, was ligated into pKGW (24). The resulting recombinant plasmids were grown in *E. coli* D1210 on M9 medium containing kanamycin and IPTG. Amplification and preparation by the method of Birnboim (3) yielded the recombinant plasmid pGM59.

Site-directed mutagenesis with an Amersham kit was performed by the method of Sayers et al. (36). Mutants were screened by *SmaI* digestion, and the mutations were confirmed by nucleotide sequence analysis.

Nucleotide sequence analysis. For nucleotide sequence analysis by the dideoxy method (35), pGM59 fragments were cloned into phages M13mp18 and M13mp19. One subclone was generated by ligating the 900-bp *Bss*HII fragment previously filled in with DNA polymerase I (Klenow fragment) (1) into the *SmaI* site of M13mp18. The resulting clone had the same orientation as *lacZ*. This recombinant was used for exonuclease III digestion (16). Both replicative-form and single-stranded phages were prepared, and the extent of the deletions was determined by agarose gel electrophoresis and DNA sequencing. Two plasmids with deletions of 150 and 350 bp were used to complete the nucleotide sequence analysis of the *Bss*HII fragment.

***aroF-lacZ* fusions.** The *aroF* regulatory regions of *E. coli* and *S. typhimurium* were subcloned into pMLB1034. The resulting recombinant plasmids were used to transform *E.*

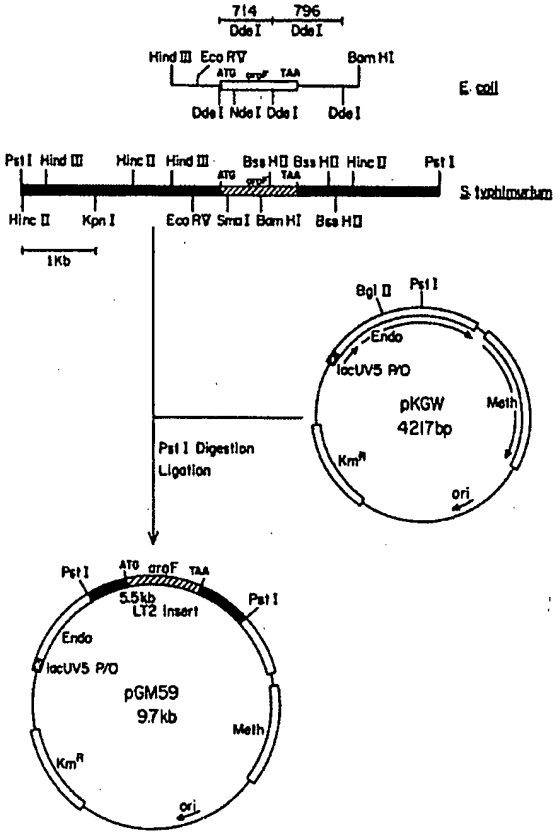


FIG. 1 Restriction endonuclease cleavage maps for the *E. coli* and *S. typhimurium aroF* genes and cloning scheme for *S. typhimurium aroF*.

coli MC4100, yielding Lac$^+$ derivatives that served as hosts to prepare lysates of λRZ5 (14) by the method of Silhavy et al. (40). Recombinant phages carrying functional *lacZ* genes were identified as blue plaques on X-Gal–agar. Six lysogens from each λRZ5 derivative carrying *aroF-lacZ* fusions were purified and shown to be single lysogens (40). The β-galactosidase activities in extracts of the lysogens were measured and are given in Miller units (29).

Purification of *S. typhimurium* tyrosine-sensitive DAHP synthase. *E. coli* GKM41 carrying plasmid pGM59 was grown to late log phase in minimal salts medium containing ampicillin. The cells were harvested by centrifugation and resuspended in 50 mM potassium phosphate (pH 6.5) containing 2% propanediol and 2 mM phosphoenolpyruvate (buffer A). The cells were disrupted in an Aminco French pressure cell at 20,000 lb/in^2; cell debris was removed by centrifugation for 60 min at 25,000 × *g*. The supernatant was treated with 0.1 volume of 2% protamine sulfate, and the precipitate was removed by centrifugation for 75 min at 30,000 × *g*. The supernatant was applied to a BioGel hydroxylapatite-HTP column equilibrated with buffer A. Protein was eluted from the column with a 500-ml linear gradient of 0.05 to 0.5 M potassium phosphate containing 2% propanediol and 2 mM phosphoenolpyruvate.

Protein was quantitated by the procedure of Lowry et al. (26); gel electrophoresis was performed on a sodium dodecyl sulfate–10% polyacrylamide gel by the method of Laemmli (25); the proteins in the gel were detected by silver staining (46). Amino acid sequencing was carried out on an Applied Biosystems model 470A gas-phase sequenator with a model 120A analyzer (17).

RESULTS

Genome blotting. *S. typhimurium* LT2 chromosomal DNA was digested separately with five restriction endonucleases. The digestion products were subjected to agarose gel electrophoresis. The dried gel was probed with *E. coli* DNA encoding tyrosine-sensitive DAHP synthase. Plasmid pCG201 (38) carries the *E. coli aroF* gene (Fig. 1) totally contained within two *Dde*I fragments of 714 and 796 bp. The autoradiogram of a gel probed with the 796-bp *Dde*I fragment is shown in Fig. 2A. The 714-bp *Dde*I probe gave an identical pattern. *Pst*I digestion yielded a 5.5-kilobase (kb) fragment that hybridized with both the 796- and 714-bp *Dde*I probes and was predicted to contain the entire *aroF* gene of *S. typhimurium*.

*Pst*I-digested *S. typhimurium* LT2 DNA was subjected to gel electrophoresis with low-melting-point agarose. Fractions containing DNA fragments of different sizes were isolated by phenol extraction of gel slices melted at 65°C. A sample of each fraction was subjected to agarose gel electrophoresis and probed as before with the 796-bp *Dde*I fragment (Fig. 2B). DNA of fraction 3 was used in the cloning experiments.

Cloning of *S. typhimurium aroF*. The positive selection vector pKGW (24) was chosen to clone *S. typhimurium aroF*. This plasmid carries the gene encoding the restriction endonuclease *Eco*RI under the control of the *lacUV5* promoter (Fig. 1). This high-copy-number plasmid is lethal to wild-type cells, but can be propagated in *lacI*^{q} cells grown in the absence of inducer. Upon induction of the *lacUV5* promoter with IPTG, overproduction of *Eco*RI becomes lethal to the host cells, unless the coding sequence of *Eco*RI is interrupted. The *Pst*I or *Bgl*II sites within the *Eco*RI coding sequence are suitable cloning sites.

Size-fractionated *Pst*I-digested *S. typhimurium* DNA was cloned in plasmid pKGW. Transformants of *E. coli* D1210 (*lacI*^{q}) were selected in the presence of IPTG and kanamycin. A library of recombinants was generated, amplified, and screened by transformation of the *aroF* strain GKM41.

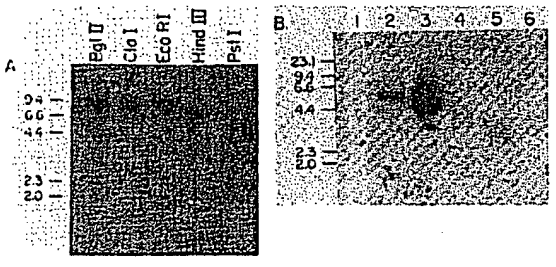


FIG. 2. Genome blotting of *S. typhimurium* DNA. (A) *S. typhimurium* DNA was digested with the indicated restriction endonucleases, and the digests were subjected to agarose gel electrophoresis. The dried gel was hybridized with the 796-bp *Dde*I probe from *E. coli* (Fig. 1). (B) *Pst*I-digested *S. typhimurium* DNA was size fractionated by agarose gel electrophoresis; samples of fractionated DNA fragments were again subjected to agarose gel electrophoresis and treated as in panel A

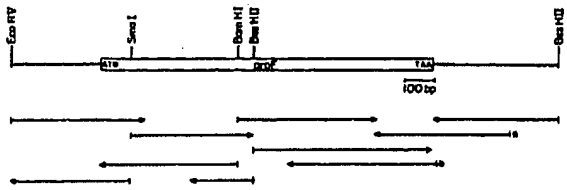


FIG. 3. Sequencing strategy for *S. typhimurium aroF*. The box indicates the coding region of *aroF*. Arrows mark the extent and direction of nucleotide sequence runs. The two clones of the exonuclease III-digested *Bss*HII fragments are marked with asterisks.

Selection was made for growth on minimal salts medium supplemented with phenylalanine and tryptophan. An *aroF*-containing plasmid designated pGM59 (Fig. 1) was isolated and shown to contain *aroF* by hybridization and DAHP synthase assay of cell extracts from plasmid-bearing strains.

Nucleotide sequence analysis of *S. typhimurium aroF*. The 5.5-kb *Pst*I insert of plasmid pGM59 was subjected to restriction analysis (Fig. 1). By hybridization analysis, *aroF* was localized within the 5.5-kb *Pst*I fragment to a 2-kb *Eco*RV-*Hin*cII fragment. Figure 3 shows the strategy that was used to obtain the nucleotide sequence of the 2-kb *Eco*RV-*Hin*cII fragment. The nucleotide sequence reveals a 1,068-bp open reading frame that is similar in sequence to *aroF* of *E. coli* (Fig. 4).

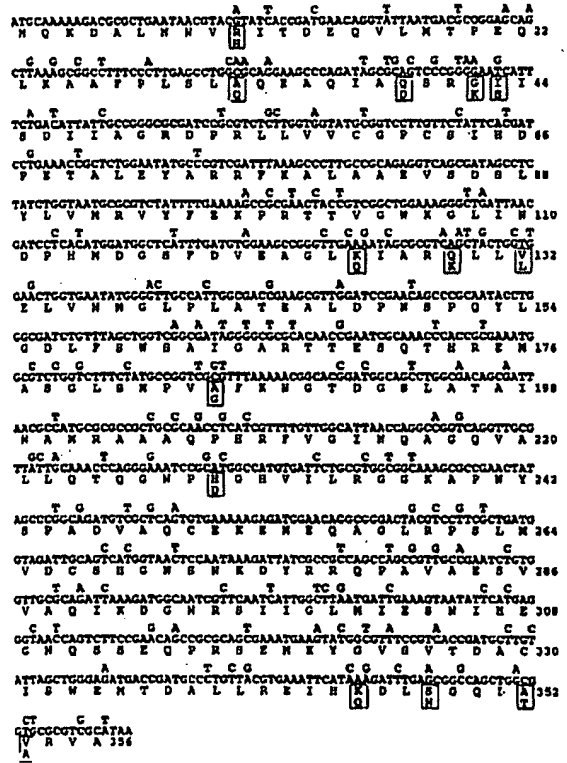


FIG. 4. Nucleotide sequence of the *S. typhimurium aroF* (GenBank no. M31302) and amino acid sequence of the encoded tyrosine-sensitive DAHP synthase (middle two lines). The sequence variations to the *E. coli aroF* coding region (upper line) and the *E. coli* DAHP synthase (lower line) are also shown.

Purification of *S. typhimurium* tyrosine-sensitive DAHP synthase and amino acid sequence analysis. The tyrosine-sensitive DAHP synthase was purified from extracts of *E. coli* GKM41 (*aroF*) carrying the *S. typhimurium aroF* plasmid pGM59. A two-step procedure yielded enzyme of near electrophoretic homogeneity (Fig. 5). The protein was subjected to N-terminal amino acid sequence analysis. The N-terminal 14 amino acid residues are identical to those predicted from the nucleotide sequence of the 1,068-bp open reading frame. Thus, we have cloned and sequenced the coding region of the *S. typhimurium aroF* gene.

Regulation of *S. typhimurium aroF* in *E. coli*. To initiate a functional analysis of the *S. typhimurium aroF* regulatory region, we constructed two very similar *lacZ* fusions to the homologous regions of the *E. coli* and the *S. typhimurium aroF* genes; the fusions contain identical coding regions. Site-directed mutagenesis was used to generate two *Sma*I sites covering codons 4 and 5 of the two genes (Fig. 6). Two M13 constructs served as templates for the mutagenesis: M13GME contains the 674-bp *Eco*RV-*Nde*I *E. coli* fragment (Fig. 1) that encompasses 334 bp upstream of the *E. coli aroF* transcription start. M13GMS contains the 434-bp *Eco*RV-*Sma*I *S. typhimurium* fragment (Fig. 1) that covers 313 bp upstream of the predicted *S. typhimurium aroF* transcription start. Both fragments were inserted into the *Sma*I site of M13mp18. After mutagenesis, new *Sma*I sites were detected by restriction analysis and confirmed by nucleotide sequence analysis. The *Eco*RI-*Sma*I fragment that contains 18 bp of the M13 polylinker 5′ to the new *E. coli Eco*RV-*Sma*I fragment of 397 bp and the *Eco*RI-*Sma*I fragment that contains the same polylinker fragment 5′ to the new *S. typhimurium Eco*RV-*Sma*I fragment of 376 bp were subcloned into the *Eco*RI and *Sma*I sites of the polylinker of plasmid pMLB1034, which carries a promoterless *lacZ* gene, yielding *aroF-lacZ* translational fusions. In the two recombinant pMLB1034 derivatives, codons 4 of the *E. coli* and the *S. typhimurium aroF* were fused to codon 6 of *lacZ*. The inserts of the recombinant plasmids were confirmed by restriction and nucleotide sequence analyses.

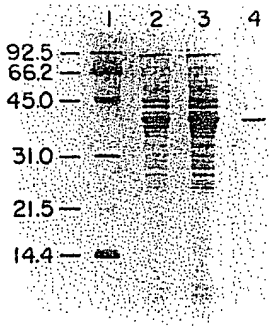
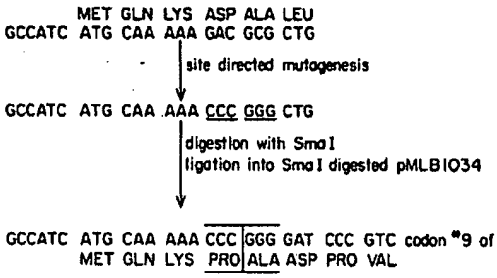


FIG. 5. Purification of *S. typhimurium* tyrosine-sensitive DAHP synthase. Protein samples were subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis, and proteins were detected by silver staining. Lanes: 1, molecular weight marker; 2, 1.4 μg of protein after treatment of the extract with protamine sulfate; 3, 2.8 μg of protein after ammonium sulfate precipitation; 4, 0.6 μg of protein from the hydroxylapatite pool



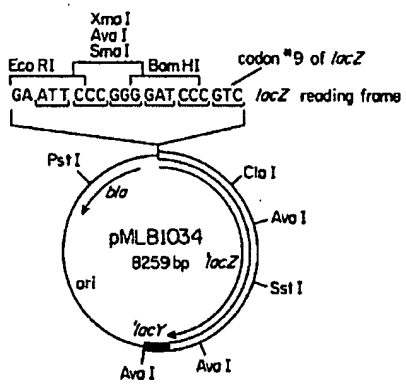


FIG. 6. Construction of *aroF-lacZ* fusions. The sequences around the site of mutagenesis are identical for *E. coli* and *S. typhimurium*. Through three single-base-pair changes a new *SmaI* site is obtained overlapping *aroF* codons 4 and 5.

Phage λRZ5 was used to place the *aroF-lacZ* fusions in single copy at the λ *att* site on the *E. coli* chromosome (14). β-Galactosidase activities in the resulting lysogens grown on media with and without tyrosine were determined (Table 2). Cells of *tyrR*$^+$ strains grown in minimal-salts tyrosine medium show very similar, low enzyme values, indicating that the *E. coli* Tyr repressor can regulate the expression of both the *E. coli* and the *S. typhimurium aroF* gene. Since the *aroF* genes differ markedly for the nucleotide sequence in the three helix turns that separate *aroFo1* and *aroFo2* (12), our results also suggest that the sequence in this region is not critical for repressor function. Under conditions of derepression, β-galactosidase activity from the *S. typhimurium aroF-lacZ* fusion is about twice as high as the activity in the corresponding strain carrying the *E. coli* fusion (Table 2).

TABLE 2. Regulation of *S. typhimurium aroF* in *E. coli*

Lysogen	200 μM tyrosine supplement	β-Galactosidase activity	
		Miller units	Relative activity
SP564(λGME4)	+	0.32 ± 0.02	1
SP564(λGME4)	−	2.3 ± 0.1	7
SP564-1(λGME4)	+	25.3 ± 1.3	80
SP564-1(λGME4)	−	26.0 ± 2.1	81
SP564(λGMS4)	+	0.3 ± 0.09	1
SP564(λGMS4)	−	4.1 ± 0.1	13
SP564-1(λGMS4)	+	49.4 ± 3.0	155
SP564-1(λGMS4)	−	46.2 ± 4.4	145

DISCUSSION

S. typhimurium chromosomal DNA contains a 5.5-kb *PstI* fragment that strongly hybridizes with *E. coli aroF* probes. The 5.5-kb *PstI* fragment was cloned into the positive selection vector pKGW. The recombinant plasmid pGM59 complements the *E. coli aroF* mutant strain GKM41. A restriction map of the 5.5-kb insert of pGM59 was obtained, and the position of *aroF* on a 2-kb fragment within the insert was determined by hybridization analysis.

The nucleotide sequence of this 2-kb fragment contains a 1,068-bp open reading frame with extensive sequence similarity to the *E. coli aroF* gene (38). The two DNA sequences are 85.1% identical. Several other genes have been sequenced from both *S. typhimurium* and *E. coli*, most notably the genes of the *trp* operon (8, 19, 30, 48). The *trp* operon genes showed between 75 and 84% sequence identity. The degree of relatedness of the two *aroF* genes is within the range at the upper end of the identity scale, perhaps reflecting the highly evolved role of DAHP synthase as a control enzyme for the common aromatic amino acid biosynthetic pathway (31).

The 85.1% identity between the two *aroF* genes contrasts with about 50% identity between *E. coli aroF* and the other two *E. coli* genes encoding DAHP synthases (9, 33, 38). The three *E. coli* genes *aroF*, *aroG*, and *aroH* are evolutionarily more distant than the homologous *S. typhimurium* and *E. coli aroF* genes, suggesting that the three DAHP synthase isoenzymes existed before the genera *Escherichia* and *Salmonella* diverged.

The *S. typhimurium aroF* gene encodes a protein of 356 amino acid residues. *S. typhimurium* tyrosine-sensitive DAHP synthase was purified, and the amino acid sequence predicted by the nucleotide sequence was confirmed through N-terminal sequence analysis. The amino acid sequences of the tyrosine-sensitive DAHP synthase from *S. typhimurium* and *E. coli* are 96.1% identical (Fig. 4). There are only 14 amino acid residue differences between the two proteins. Eleven of these differences are localized to three small regions of the protein, residues 32 to 43, 125 to 132, and 345 to 353. Within these regions the sequence similarity is only 55.6, 66.3, and 66.7%, respectively. Apparently, none of these differences have a marked influence on the kinetic or stability properties of the enzyme. The apparent V_{max} values for both enzymes are very similar, and both enzymes are subject to rapid inactivation when cells approach the stationary phase of growth (G. K. Muday and K. M. Herrmann, unpublished data), a phenomenon previously described for the *E. coli* enzyme (43).

The codon usage of the two *aroF* genes is relatively close to the overall codon usage of *E. coli*, as well as to each other. It is, however, quite interesting that in both *aroF* genes rare codons are used for the aromatic amino acids tyrosine and phenylalanine. In *E. coli*, 40% of all tyrosine codons are TAT; for the *aroF* genes, five of seven *E. coli* and six of seven *S. typhimurium* codons are TAT. Likewise in *E. coli*, 37% of all phenylalanine codons are TTT; for the *aroF* genes of both organisms, seven of seven codons are TTT. Also, several other rare codons are used slightly more frequently in the *S. typhimurium* gene, for example, TTA and TTG for

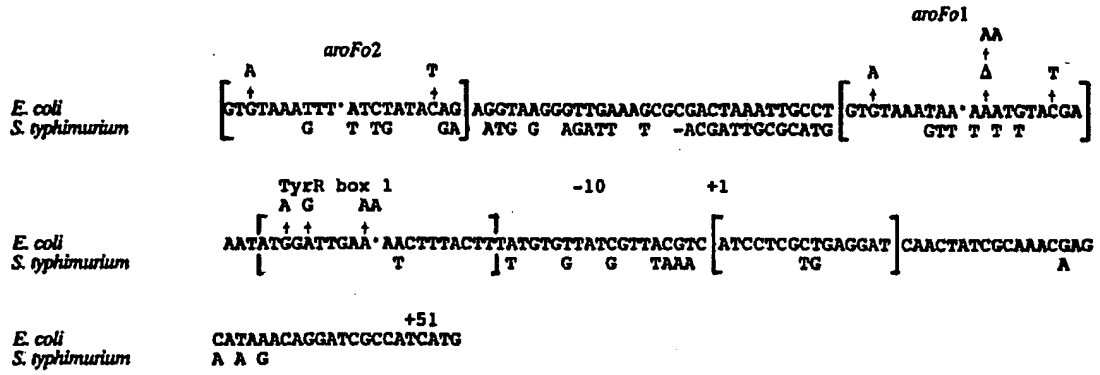


FIG. 7. Comparison of *aroF* regulatory regions of *E. coli* and *S. typhimurium*. The complete *E. coli* sequence and the *S. typhimurium* variations are given. Brackets designate operator boxes *aroFo1*, *aroFo2*, TyrR box 1, and a palindromic sequence of unknown significance. Identified operator mutations (7, 12) are indicated by the arrows above the sequences.

leucine, ATA for isoleucine, and CGA and CGG for arginine.

A more quantitative analysis showed a correlation of optimal codon usage to cellular levels of the protein products (21). When this analysis is applied to the two *aroF* genes, the frequencies are 0.61 and 0.64 for *S. typhimurium* and *E. coli*, respectively. The *E. coli aroF* value is close to 0.63, 0.65, and 0.66 for the *E. coli trpA*, *trpD*, and *trpE* genes, respectively. The frequencies of optimal codon usage for *S. typhimurium trpA*, *trpB*, and *trpD* are all 0.63, slightly lower than the *E. coli trp* gene values, just as the *S. typhimurium aroF* gene has a lower value than the *E. coli aroF* gene. This result is in good agreement with a few other examples (13) that show higher frequencies of optimal codon usage in *E. coli*, suggestive of a slightly tighter codon preference.

The regulatory regions of the two *aroF* genes were also compared. For the *E. coli* gene, the start of transcription has been determined (12), a strong promoter has been identified (12), and three operator boxes have been defined by mutational analysis (7, 12). The *aroF* regulatory region of *S. typhimurium* shows a high degree of sequence identity to the corresponding *E. coli* sequences of defined regulatory function, but very little similarity in the sequences outside the defined regions (Fig. 7). In all *E. coli aroF* operator mutations identified thus far, the changes are in nucleotides that are identical for both wild-type organisms (Fig. 7). The *S. typhimurium* promoter was assigned on the basis of its similarity in sequence and location to the *E. coli* promoter. The promoters are very similar and very strong, with McClure algorithm scores of 68.6 and 69.2 for the *S. typhimurium* and the *E. coli* promoter, respectively (15).

Three operator boxes were identified within the *S. typhimurium* regulatory region, based on their similarity to the *E. coli* boxes. The sequence and location of the boxes and the distance between them are very similar for both genes (Fig. 7). The distance between *aroFo1* and *aroFo2* corresponds to three turns of the helix. Although the distance between these operator boxes is preserved, the nucleotide sequence in the region between the boxes shows much less similarity, indicating that the distance, if not the sequences, between the boxes may be important for proper operator recognition by the *tyrR* product.

The high degree of similarity between the regulatory regions of the two *aroF* genes suggests that the levels of transcription and translation of these two genes should be very similar. To compare the control features for both *aroF* genes, we fused their regulatory regions to *lacZ* and measured the β-galactosidase levels in strains carrying the translational fusions in single copy. Under derepression conditions, expression of the *S. typhimurium* promoter is about twice that of the *E. coli* promoter. This difference could account in part for the higher levels of the tyrosine-sensitive DAHP synthase reported for wild-type *S. typhimurium* (22). Under repression conditions, the enzyme values were about the same. These results indicate that the *E. coli* Tyr repressor recognizes both *E. coli* and *S. typhimurium aroF* operators with about equal efficiency. Since the sequences that separate *aroFo1* and *aroFo2* in the two organisms show no similarity, one may assume that the sequences within these three turns of the DNA helix cannot be part of the recognition signal for repressor binding. A preliminary study with a multicopy plasmid system suggested that the distance between *E. coli aroFo1* and *aroFo2* is also not critical for repressor recognition (6). However, the separation of the two operator boxes by three turns of a helix for both organisms suggests an important role for this portion of the regulatory region. The regulatory region of the *E. coli aroL* gene, which is also controlled by the tyrosine repressor, also has three operator boxes, arranged in the same way as the *aroF* boxes (10). The distance between the two upstream *aroL* boxes is again three turns of the helix, yet the nucleotide sequence of the spacer shows no similarity to the corresponding *aroF E. coli* or *S. typhimurium* sequences.

A new potential regulatory region is also seen when the *E. coli* and *S. typhimurium aroF* gene sequences are compared. In the untranslated portion of the transcript for both regulatory regions, a palindromic sequence is found with a potential for stem-loop formation. The regulatory significance of these palindromes will be addressed through site-directed mutagenesis.

ACKNOWLEDGMENTS

We thank R. L. Somerville and H. Zalkin for bacterial strains and plasmids, M. A. Hermodson for help with the amino acid sequence analysis, and R. L. Somerville and B. L. Wanner for critical reading of the manuscript.

G.K.M. was supported by David Ross grant 6901264.

LITERATURE CITED

1. Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Siedman, J. A. Smith, K. Struhl. 1987. Current protocols in molecular biology. Greene Publishing Associates and John Wiley & Sons, Inc., New York, N.Y.

2. Beltz, G. A., K. A. Jacobs, T. H. Eickbush, P. T. Cherbas, and F. C. Kafatos. 1983. Isolation of multigene families and determination of homologies by filter hybridization methods. Methods Enzymol. 100:266–285.
3. Birnboim, H. C. 1983. A rapid alkaline extraction method for the isolation of plasmid DNA. Methods Enzymol. 100:243–249.
4. Bochner, B. R., H. C. Huang, G. L. Schieven, and B. N. Ames. 1980. Positive selection for loss of tetracycline resistance. J. Bacteriol. 143:926–933.
5. Clark, V. M., and A. J. Kirby. 1963. An improved synthesis of phosphoenolpyruvate. Biochim. Biophys. Acta 78:732.
6. Cobbett, C. S. 1988. Repression of the *aroF* promoter by the Tyr repressor in *Escherichia coli* K-12: role of the upstream operator site. Mol. Microbiol. 2:377–383.
7. Cobbett, C. S., and M. L. Delbridge. 1987. Regulatory mutants of the *aroF-tyrA* operon of *Escherichia coli* K-12. J. Bacteriol. 169:2500–2506.
8. Crawford, I. P., B. P. Nichols, and C. Yanofsky. 1980. Nucleotide sequence of the *trpB* gene in *Escherichia coli* and *Salmonella typhimurium*. J. Mol. Biol. 142:489–502.
9. Davies, W. D., and B. E. Davidson. 1982. The nucleotide sequence of *aroG*, the gene for 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthetase (phe) in *Escherichia coli* K-12. Nucleic Acids Res. 10:4045–4058.
10. DeFeyter, R. C., B. E. Davidson, and J. Pittard. 1986. Nucleotide sequence of the transcription unit containing the *aroL* and *aroM* genes from *Escherichia coli* K-12. J. Bacteriol. 165: 233–239.
11. Doy, C. H., and K. D. Brown. 1965. Control of aromatic biosynthesis: the multiplicity of 7-phospho-2-oxo-3-deoxy D-*arabino*-heptonate D-erythrose-4-phosphate-lyase (pyruvate-phosphorylating) in *Escherichia coli* W. Biochim. Biophys. Acta 104:377–389.
12. Garner, C. C., and K. M. Herrmann. 1985. Operator mutations of the *Escherichia coli aroF* gene. J. Biol. Chem. 260:3820–3825.
13. Gouy, M., and C. Gautier. 1982. Codon usage in bacteria: correlation with gene expressivity. Nucleic Acids Res. 10:7055–7074.
14. Grove, C. L., and R. P. Gunsalus. 1987. Regulation of the *aroH* operon of *Escherichia coli* by the tryptophan repressor. J. Bacteriol. 169:2158–2164.
15. Hawley, D. K., and W. R. McClure. 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res. 11:2237–2255.
16. Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene 28: 351–359.
17. Hermodson, M. A., L. H. Ericsson, K. Titani, H. Neurath, and K. A. Walsh. 1972. Application of sequenator analyses to the study of proteins. Biochemistry 11:4493–4502.
18. Herrmann, K. M. 1983. The common aromatic biosynthetic pathway, p. 301–322. *In* K. M. Herrmann and R. L. Somerville (ed.), Amino acids: biosynthesis and genetic regulation. Addison-Wesley Publishing Co., Reading, Mass.
19. Horowitz, H., J. van Arsdell, and T. Platt. 1983. Nucleotide sequence of the *trpD* and *trpC* genes of *Salmonella typhimurium*. J. Mol. Biol. 169:775–797.
20. Hu, C.-Y., and D. B. Sprinson. 1977. Properties of tyrosine-inhibitable 3-deoxy-D-*arabino*-heptulosonic acid-7-phosphate synthase from *Salmonella*. J. Bacteriol. 129:177–183.
21. Ikemura, T., and H. Ozeki. 1982. Codon usage and transfer RNA contents: organism-specific codon choice patterns in reference to the isoacceptor contents. Cold Spring Harbor Symp. Quant. Biol. 47:1087–1097.
22. Jensen, R. A., D. S. Nasser, and E. W. Nester. 1967. Comparative control of a branchpoint enzyme in microorganisms. J. Bacteriol. 94:1582–1593.
23. Kleckner, N., J. Roth, and D. Botstein. 1977. Genetic engineering *in vivo* using translocatable drug-resistance elements. J. Mol. Biol. 116:125–159.
24. Kuhn, I., F. H. Stephenson, H. W. Boyer, and P. J. Greene. 1986. Positive selection vectors utilizing lethality of the *Eco*RI endonuclease. Gene 42:253–263.
25. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) 227:680–685.
26. Lowry, O. H., N. J. Rosebrough, A. L. Farr, and R. J. Randall. 1951. Protein measurements with the Folin phenol reagent. J. Biol. Chem. 193:265–275.
27. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
28. Messing, J. 1983. New M13 vectors for cloning. Methods Enzymol. 101:20–78.
29. Miller, J. H. 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
30. Nichols, B. P., and C. Yanofsky. 1979. Nucleotide sequence of *trpA* of *Salmonella typhimurium* and *Escherichia coli*: an evolutionary comparison. Proc. Natl. Acad. Sci. USA 76:5244–5248.
31. Ogino, T., C. Garner, J. L. Markley, and K. M. Herrmann. 1982. Biosynthesis of aromatic compounds: ^{13}C NMR spectroscopy of whole *Escherichia coli* cells. Proc. Natl. Acad. Sci. USA 79:5828–5832.
32. Pittard, J. 1987. Biosynthesis of the aromatic amino acids, p. 368–394. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
33. Ray, J. M., C. Yanofsky, and R. Bauerle. 1988. Mutational analysis of the catalytic and feedback sites of the tryptophan-sensitive 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase of *Escherichia coli*. J. Bacteriol. 170:5500–5506.
34. Saito, H., and K. I. Miura. 1963. Preparation of transforming deoxyribonucleic acid by phenol treatment. Biochim. Biophys. Acta 72:619–629.
35. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463–5467.
36. Sayers, J. R., W. Schmidt, and F. Eckstein. 1988. 5'-3'Exonucleases in phosphothionate-based oligonucleotide-directed mutagenesis. Nucleic Acids Res. 16:791–802.
37. Schoner, R., and K. M. Herrmann. 1976. 3-Deoxy-D-*arabino*-heptulosonate 7-phosphate synthase. Purification, properties and kinetics of the tyrosine-sensitive isoenzyme from *Escherichia coli*. J. Biol. Chem. 251:5440–5447.
38. Shultz, J., M. A. Hermodson, C. C. Garner, and K. M. Herrmann. 1984. The nucleotide sequence of the *aroF* gene of *Escherichia coli* and the amino acid sequence of the encoded protein, the tyrosine-sensitive 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase. J. Biol. Chem. 259:9655–9661.
39. Sieben, A. S., A. S. Perlin, and F. J. Simpson. 1966. An improved preparative method for D-erythrose-4-phosphate. Can. J. Biochem. 44:663–669.
40. Silhavy, T. J., M. L. Berman, and L. W. Enquist. 1984. Experiments with gene fusions. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
41. Srinivasan, P. R., and D. B. Sprinson. 1959. 2-Keto-3-deoxy-D-*arabo*-heptonic acid 7-phosphate synthetase. J. Biol. Chem. 234:716–722.
42. Tribe, D. E., H. Camakaris, and J. Pittard. 1976. Constitutive and repressible enzymes of the common pathway of aromatic biosynthesis in *Escherichia coli* K-12: regulation of enzyme synthesis at different growth rates. J. Bacteriol. 127:1085–1097.
43. Tribe, D. E., and J. Pittard. 1979. Hyperproduction of tryptophan by *Escherichia coli*: Genetic manipulation of the pathways leading to tryptophan formation. Appl. Environ. Microbiol. 38:181–190.
44. Tsao, S. G. S., C. F. Brunk, and R. E. Pearlman. 1983. Hybridization of nucleic acids directly in agarose gels. Anal. Biochem. 131:365–372.
45. Vogel, H. J., and D. M. Bonner. 1956. Acetylornithinase of

Escherichia coli: partial purification and some properties. J Biol Chem 218:97–106.
46. Wray, W., R. Boulikas, V. P. Wray, and R. Hancock. 1981. Silver staining of proteins in polyacrylamide gels. Anal. Biochem. 118:197–203.
47. Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103–119.
48. Yanofsky, C., and M. vanCleemput. 1982. Nucleotide sequence of *trpE* of *Salmonella typhimurium* and its homology with the corresponding sequence of *Escherichia coli*. J. Mol. Biol. 155: 235–246.